Building Models of Protein Complexes using the Known Protein-protein Interfaces: Structural Model of PhoB Dimeric Complex in its Active State

Chang-Shung Tung, T-6

Often, multimolecular complexes are the functional forms of proteins inside cells. To understand how proteins function, it is necessary to know how they interact with each other. Knowing the structures of protein complexes is critical. It is well known that solving the structure of complexes is significantly more difficult than solving the structure of individual proteins. While homology modeling is a promising approach in modeling protein structures, it has not been proved successful in modeling protein complex structures. Improving the knowledge of protein complex structures remains a great challenge. The good news is that a wealth of structural information exists. In the current version of Protein Data Bank (PDB) it can be seen that many of these proteins exist in a multimer conformation that contains the information of interaction surfaces (interfaces). Here we present a novel approach that utilizes the information of known protein interfaces to build models of protein complexes. This approach will significantly improve protein complex structural modeling and prediction.

The protein complex that is used to illustrate the approach is the PhoB dimer in its functional active state, which binds to the targeted DNA duplex. In the crystal structure (PDB accession code: 1GXP), two PhoB EDs (magenta and white molecules) bind to the targeted DNA direct repeat pho box (cyan and yellow molecules) as shown in Fig.1a. The binding of DNA direct repeats makes the ED dimer follow a tandem symmetry. However, the known structure of the PhoB RD dimeric complex (PDB accession code: 2JB9) follows a rotational symmetry. It is not obvious how to best dock a RD dimeric complex with a rotational symmetry onto an ED dimeric complex with a tandem symmetry. A number of response regulators exist with their two-domain structures solved experimentally (PDB accession codes: 1KGS, 1P2F, 1YS6, 2GWR, 2OQR, 1A04, 1YIO). These structures contain the information of ED/RD interfaces. Applying the information together with the structure of the ED/ED dimeric complex (1GXP), we explore the potential solutions for the PhoB dimeric complex. Out of the ED/RD conformations, only that of DrrB (1P2F , shown as the red and the blue molecules in Fig. 1), a PhoB/OmpR homolog, provides a satisfactory solution where the two RDs are in contact but not overlapping. Combining the structural information of ED/ED (1GXP), ED/RD (1P2F), ED (1GXP) and RD (2JB9), the model of the PhoB dimeric complex is developed (shown as the white/magenta molecules bound to DNA in Fig. 1b). This model structure has appealing features including: good stereochemistry (no clashes between domains, stable interface surface area), protein-like structure (contents of secondary structures, density, etc.), and several of the known interaction interfaces. However, the RD/RD dimer in the modeled complex follows a tandem symmetry that is different from the rotational symmetry observed in the PhoB RD/RD dimer (2JB9). The remaining question is: Does the new interface between the two RDs in the model exist in other protein domains of a similar fold? To answer this question, we search for interfaces between domains that have the flavodoxin-like fold and give the two domains a tandem symmetry. Interestingly, the dimeric structure (PDB accession code: 1FFY) of CheY, a chemotaxis protein, has two flavodoxin-like domains following a tandem symmetry. This dimeric structure of CheY (1FFY) is very similar to that of the PhoB dimeric RDs as shown in Fig.1c. We have demonstrated that the information of protein interfaces and folds can be directly used to build structural models of protein complexes.

For more information contact Chang-Shung Tung at ct@lanl.gov.

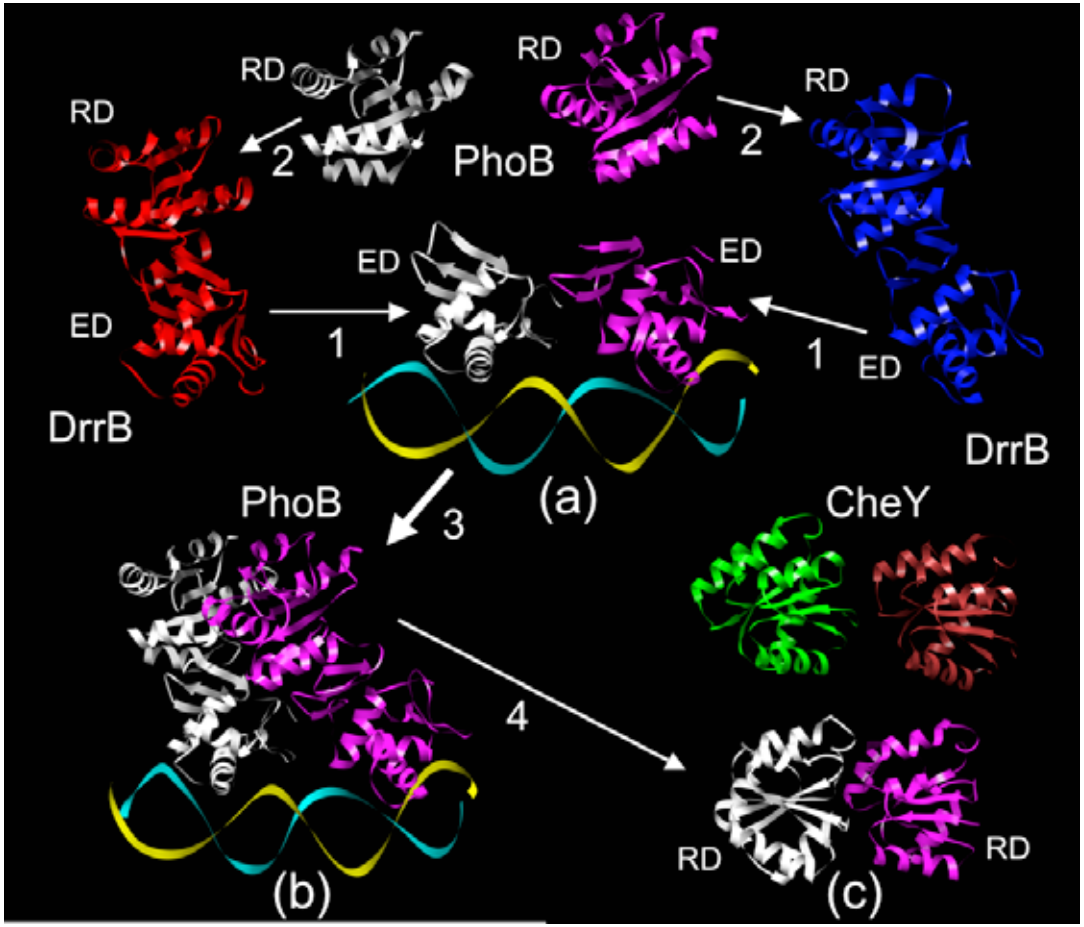

Fig. 1. Structural model of the PhoB dimeric complex as it binds to its targeted DNA duplex. The matching ED of DrrB to ED of PhoB is shown in1a. The resultant PhoB/DNA model that is shown in 1b. The RD/RD interface in CheY is similar to that in the model as shown in 1c.

Funding Acknowledgments
LANL Directed Research and Development Program